



Kalman filters for assimilating near-surface observations into the Richards equation – Part 1: Retrieving state profiles with linear and nonlinear numerical schemes

G. B. Chirico¹, H. Medina², and N. Romano¹

¹Department of Agricultural Engineering, University of Naples Federico II, Naples, Italy

²Department of Basic Sciences, Agrarian University of Havana, Havana, Cuba

Correspondence to: G. B. Chirico (gchirico@unina.it)

Received: 3 November 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 3 December 2012

Revised: 10 January 2014 – Accepted: 14 January 2014 – Published: 4 July 2014

Abstract. This paper examines the potential of different algorithms, based on the Kalman filtering approach, for assimilating near-surface observations into a one-dimensional Richards equation governing soil water flow in soil. Our specific objectives are: (i) to compare the efficiency of different Kalman filter algorithms in retrieving matric pressure head profiles when they are implemented with different numerical schemes of the Richards equation; (ii) to evaluate the performance of these algorithms when nonlinearities arise from the nonlinearity of the observation equation, i.e. when surface soil water content observations are assimilated to retrieve matric pressure head values. The study is based on a synthetic simulation of an evaporation process from a homogeneous soil column. Our first objective is achieved by implementing a Standard Kalman Filter (SKF) algorithm with both an explicit finite difference scheme (EX) and a Crank-Nicolson (CN) linear finite difference scheme of the Richards equation. The Unscented (UKF) and Ensemble Kalman Filters (EnKF) are applied to handle the nonlinearity of a backward Euler finite difference scheme. To accomplish the second objective, an analogous framework is applied, with the exception of replacing SKF with the Extended Kalman Filter (EKF) in combination with a CN numerical scheme, so as to handle the nonlinearity of the observation equation. While the EX scheme is computationally too inefficient to be implemented in an operational assimilation scheme, the retrieval algorithm implemented with a CN scheme is found to be computationally more feasible and accurate than those implemented with the backward Euler scheme, at least for the examined one-dimensional problem. The UKF appears to be as feasible as the EnKF when one has to handle non-

linear numerical schemes or additional nonlinearities arising from the observation equation, at least for systems of small dimensionality as the one examined in this study.

1 Introduction

Soil water in the vadose zone exerts a large control on the water and energy balance of land-atmosphere systems over a wide range of space-time scales (e.g. Milly and Dunne, 1994; Entekhabi et al., 1996; Vrugt et al., 2001, 2003; Rodriguez-Iturbe and Porporato, 2005). With the increasing availability of near-surface data from remote and ground-based sensors, unique opportunities emerge to predict the soil water dynamics (McLaughlin, 2002; Vereecken et al., 2008). A key challenge is to identify the best approaches for efficiently integrating these data with the soil water dynamic models, in order to achieve more reliable and purposeful predictions. Hence, data assimilation has become a relatively important area of investigation, aiming at an efficient integration of remote-sensing techniques, ground-based sensors and soil water dynamic models (Hoebein and Troch, 2000; Heathman et al., 2003; de Lannoy et al., 2007; Matgen et al., 2010).

The physics of isothermal flow in unsaturated soils is commonly modelled with the Richards equation (Jury et al., 1991). Three standard forms of the unsaturated flow equation can be found in literature: (i) the “ h -based form” and (ii) the “ θ -based form”, whether the dependent variable is matric pressure head, h [L], or soil water content θ [L³ L⁻³], respectively; (iii) the “mixed form” when both the dependent

variables are employed. The water retention $\theta(h)$ and the hydraulic conductivity $K(\theta)$ [L T^{-1}] functions provide constitutive relationships between those two variables and the hydraulic conductivity K , allowing for conversion of one form of the equation to the other one.

A primary source of numerical difficulty when dealing with the Richards equation is its strongly nonlinear nature. The standard numerical approximations that are applied to the spatial domain are the finite difference method and the finite element method. For any Euler method other than the fully explicit forward method, nonlinear algebraic equations result and some linearization and/or iteration procedure must be implemented to solve the discrete equations (Celia et al., 1990).

The Kalman Filter (Kalman, 1960) is a sequential data assimilation technique, largely employed in hydrological applications to merge observations arising from different sources, such as remote-sensing instruments or ground-based sensors, with dynamic models (McLaughlin, 2002), and it can be considered as the most general estimator for linear dynamic systems (Vereecken et al., 2008). The Kalman Filter “optimally” weights the a priori model state predictions at a given time with the measurements available at the same time, according to a least squares approach, in order to provide a posteriori estimates of the state system evolution.

Although the standard Kalman Filter (SKF) was originally formulated for an optimal estimation of linear state space models with Gaussian uncertainties, more KF algorithms have been developed to handle nonlinear models. The Extended Kalman Filter (EKF), which relies on the linearization of model using first order approximation of Taylor series, was the first variant designed for dealing with nonlinear models. Katul et al. (1993) and Entekhabi et al. (1994) applied the EKF for the estimation of the vertical soil moisture profile with nonlinear numerical schemes of the Richards equation. Walker et al. (2001) compared SKF and direct insertion assimilation schemes within a synthetic study similar to that adopted by Entekhabi et al. (1994), but employing a linear explicit finite difference scheme of the Richards equation.

Further developments of the Kalman Filter have been suggested to overcome drawbacks of the EKF, which have been reported in case of strong nonlinearities and high dimensional applications (e.g. Reichle et al., 2002b; van der Merwe, 2004).

Evensen (1994) proposed the Ensemble Kalman Filter (EnKF), based upon Monte Carlo generations of an ensemble of states to approximate the propagation of the state prediction error statistics through the nonlinear model. The EnKF is relatively easy to implement and it can efficiently deal with high dimensional applications. It has become a popular choice for data assimilation in traditional hydrological applications like the estimation of streamflow (Moradkhani et al., 2005; Clark et al., 2008; Weerts and El Serafy, 2006; Xie and Zhang, 2010), land surface energy fluxes (Dunne and Entekhabi, 2006; Pipunic et al., 2008) and soil moisture (Re-

ichle and Koster, 2003; Reichle et al., 2007; De Lannoy et al., 2007). It has also been applied in subsurface models based on the numerical solver of the Richards equation (Das and Mohanty, 2006; Huang et al., 2007; Camporese et al., 2009).

A method less commonly applied in hydrological studies is the Unscented Kalman Filter (UKF) developed by Julier et al. (1995) and Julier and Uhlmann (1997, 2004), also based on propagating an ensemble of sample states, but chosen in a deterministic way. Compared with the EnKF, the UKF is also expected to be less computationally efficient for systems of large dimensionality and its implementation is less straightforward. The results of Luo and Moroz (2009, 2010) suggest that in small scale applications the UKF could perform slightly better than the EnKF in response to the relatively large biases and spurious modes of Monte Carlo approximations. To our knowledge, the UKF has not been implemented yet for retrieving soil water state profiles with a soil water transport model based on a numerical solution of the Richards equation.

The assimilation of near-surface information into the Richards equation can be treated by two alternative approaches: (i) using a Standard Kalman Filter (SKF), providing an optimal estimate of the mean and error variance of the state variable for a linear numerical solver of the Richards equation, or (ii) using a non-standard KF, such as the EKF, the UKF and the EnKF, which supplies an approximate solution of the first two moments of the state variable, but with a nonlinear numerical scheme. Examining the advantages and limitations of these two alternative approaches can be relevant to identify the best strategy for implementing assimilation algorithms in operational soil hydrological studies. Nonlinear numerical schemes outperform linear schemes, since they allow us to achieve numerical stability and accuracy with much larger time-steps than linear schemes (e.g. Haverkamp et al., 1977; Paniconi et al., 1991). However, the larger computational effort required for the application of linear schemes might be compensated by applying SKF analytic estimators for the propagation of the first two moments of the state profiles, with minor computational costs than the corresponding non-standard Kalman Filter algorithms.

We are not aware of previous studies that have explicitly examined the relative efficiency of these alternative approaches.

The general aim of this paper is to compare the efficiency of soil water state profile retrieval algorithms involving standard and non-standard Kalman Filter methods applied respectively to linear and nonlinear numerical schemes of the Richards equation. These analyses are conducted by repeating the same synthetic experiment implemented by Entekhabi et al. (1994) and Walker et al. (2001), simulating an evaporation process from a homogeneous soil column.

Two scenarios are examined in this paper. The first concerns the assimilation of surface matric head for retrieving matric head profiles with h -based numerical schemes of the Richards equation. Under this scenario, the following retrieval algorithms are compared: SKF applied to both the explicit and Crank–Nicolson linear numerical schemes; the UKF and EnKF applied to a nonlinear numerical scheme.

The second scenario differs from the first as the assimilated surface variable is the soil water content instead of the matric head. In this case, in order to handle the nonlinearity arising from the discrepancy between the observed and the retrieved variables, the SKF is replaced by the EKF in combination with the Crank–Nicolson linear numerical scheme, while the explicit numerical scheme is excluded from the comparison study.

The paper is structured as follows: Sect. 2 illustrates the Kalman Filter algorithms employed in this study; Sect. 3 presents the different numerical schemes of the Richards equation; Sect. 4 describes the results of the numerical experiments; Sect. 5 is devoted to the conclusions.

2 Kalman filtering

The Kalman Filter is a recursive filter that estimates the state of a dynamic system from a series of noise corrupted measurements. Its basic theory has originally been designed for linear systems (Kalman, 1960), but several variants have subsequently been proposed for studying the dynamics of nonlinear systems.

In the most general case, the dynamic system and the measurements are described by two sets of equations, discretised in the time domain (e.g. van der Merwe; 2004):

$$\mathbf{x}_k = F_{k-1,k}(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{v}_{k-1} \quad (1)$$

$$\mathbf{y}_k = H_k(\mathbf{x}_k) + \boldsymbol{\eta}_k \quad (2)$$

where $F_{k-1,k}$ is the dynamic system model that propagates the state vector \mathbf{x} in time, assuming discrete time steps k ; \mathbf{u}_k represents the current exogenous input vector, which is assumed to be known; H_k is the measurement model, which describes how the current measurements vector \mathbf{y}_k is related to the current state \mathbf{x}_k . The dynamic system model is assumed to be corrupted by a zero mean additive Gaussian noise \mathbf{v}_{k-1} with covariance \mathbf{Q}_{k-1} . Similarly, the measurement model is assumed to be corrupted by a zero mean additive Gaussian noise vector $\boldsymbol{\eta}_k$ with covariance \mathbf{R}_k .

With respect to the more general Bayesian theory, the system state \mathbf{x}_k evolves over time according to a hidden Markov process, with a conditional probability density $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ fully specified by $F_{k-1,k}$ and by the process noise distribution $p(\mathbf{v}_{k-1})$. The observations \mathbf{y}_k are conditionally independent given the state and are generated according to the conditional probability density $p(\mathbf{y}_k|\mathbf{x}_k)$, which is fully specified by H_k and the observation noise distribution $p(\boldsymbol{\eta}_k)$.

The Kalman Filter provides a posteriori estimates of the first two moments of the state distribution:

- the mean state $\hat{\mathbf{x}}_k = E[\mathbf{x}_k]$, corresponding to the estimated state;
- the covariance of the state distribution $\mathbf{P}_k = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T]$, which is equivalent to the error covariance matrix, i.e. a measure of the accuracy of the estimated state.

The two moments are computed according to two different phases: a prediction phase and an update phase. During the prediction phase, an a priori estimate of the state $\hat{\mathbf{x}}_k^-$ and its covariance matrix \mathbf{P}_k^- are provided based on the information available at time step t_{k-1} .

The update phase is activated as the measurements \mathbf{y}_k become available. In this phase, an a posteriori state estimate $\hat{\mathbf{x}}_k$ is provided by a linear combination of the a priori estimate $\hat{\mathbf{x}}_k^-$ and the measurement innovation vector, equal to the difference between the actual measurements \mathbf{y}_k and the a priori prediction of the measurements $\hat{\mathbf{y}}_k^-$:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k^-). \quad (3)$$

In Eq. (3), the innovation vector is weighted through the matrix \mathbf{K}_k , expressed as a function of the cross covariance matrix of the state prediction error and the observation prediction error $\mathbf{P}_{xy,k}$, and the auto-covariance matrix of the predicted measurement $\mathbf{P}_{yy,k}$:

$$\mathbf{K}_k = \mathbf{P}_{xy,k} [\mathbf{P}_{yy,k} + \mathbf{R}_k]^{-1}. \quad (4)$$

The a posteriori error covariance \mathbf{P}_k is estimated as follows:

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k [\mathbf{P}_{yy,k} + \mathbf{R}_k] \mathbf{K}_k^T. \quad (5)$$

Equation (5) represents an expression of the system covariance alternative to the common one $\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_k^-$ (Julier and Uhlmann, 2004; van der Merwe, 2004). This last relationship explicitly presumes that the measurement operator H_k is linear, hence represented by a matrix \mathbf{H}_k . We prefer using the more general expression as in Eq. (5), since we also examine the case of nonlinear measurement operators in this study.

2.1 Standard Kalman Filter (SKF) and Extended Kalman Filter (EKF)

The SKF involves linear dynamic system and measurement operators and thus the operator $F_{k-1,k}$ and H_k are described by matrices, indicated as $\mathbf{F}_{k-1,k}$ and \mathbf{H}_k , respectively. The a posteriori state $\hat{\mathbf{x}}_k$ is the optimal estimate in this case, with the minimum mean square error. The system covariance \mathbf{P}_k^- , the cross-covariance matrix between the error in $\hat{\mathbf{x}}_k^-$ and the error in $\hat{\mathbf{y}}_k^-$, $\mathbf{P}_{xy,k}$, and the covariance of the predicted measurements, $\mathbf{P}_{yy,k}$, can be computed by the following closed linear relations:

$$\mathbf{P}_k^- = \mathbf{F}_{k-1,k} \mathbf{P}_{k-1} \mathbf{F}_{k-1,k}^T + \mathbf{Q}_{k-1} \quad (6)$$

$$\mathbf{P}_{xy,k} = \mathbf{P}_k^- \mathbf{H}_k^T \quad (7)$$

$$\mathbf{P}_{yy,k} = \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T. \quad (8)$$

The Kalman gain can be then computed with Eq. (4), the a posteriori state estimate $\hat{\mathbf{x}}_k$ with Eq. (3) and the a posteriori error covariance \mathbf{P}_k with Eq. (5).

If the observation operator H_k in Eq. (2) is nonlinear (a scenario analysed in this work), SKF is not applicable and a different algorithm has to be adopted to linearize H_k . If one adopts the EKF linearizing strategy, the Kalman gain and the a posteriori estimate of the covariance matrix take the following forms:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{C}_k^T \left(\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{D}_k \mathbf{R}_k \mathbf{D}_k^T \right)^{-1} \quad (9)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^- \quad (10)$$

where $\mathbf{C}_k = \partial H(\mathbf{x}, \boldsymbol{\eta}_k) / \partial \mathbf{x}|_{\hat{\mathbf{x}}_k^-}$ is the Jacobian matrix of H_k with respect to the state vector \mathbf{x} computed at the a priori estimate $\hat{\mathbf{x}}_k^-$, and $\mathbf{D}_k = \partial H(\hat{\mathbf{x}}_k^-, \boldsymbol{\eta}) / \partial \boldsymbol{\eta}|_{\bar{\boldsymbol{\eta}}=0}$ is the Jacobian matrix of H_k with respect to the noise vector $\boldsymbol{\eta}$ at the mean value $\bar{\boldsymbol{\eta}}=0$.

2.2 Unscented Kalman Filter (UKF)

The Unscented Kalman Filter (UKF) belongs to a wider group of approaches known as Sigma Point Kalman Filters (van der Merwe, 2004). The UKF is based on the Unscented Transformation, originally introduced by Julier and Uhlman (1997, 2004) as an effective method for capturing the nonlinear propagation of the first two moments of the state distribution through a minimal set of deterministically chosen sample points.

For a state vector of dimension N , the UKF, in its basic mode, obtains a set of sigma points, consisting of $2N+1$ vectors and their associated weights, $\mathbf{S} = \{\mathcal{X}_i, \mu_i^{(j)}; i = 0 \dots 2N; j \in (m, c)\}$, completely capturing the actual mean and covariance of the random variable \mathbf{x} . The vectors are weighted according to the respective mean (m) and covariance (c) weights $\mu_i^{(j)}$. A selection of sigma points fulfilling this requirement is defined as follows:

$$\begin{aligned} \mathcal{X}_0 &= \hat{\mathbf{x}}; \mathcal{X}_i = \hat{\mathbf{x}} + \left(\sqrt{\gamma \mathbf{P}} \right)_i, i = 1, \dots, N; \\ \mathcal{X}_i &= \hat{\mathbf{x}} - \left(\sqrt{\gamma \mathbf{P}} \right)_i, i = N+1, \dots, 2N \\ \mu_0^{(m)} &= \frac{\gamma - N}{\gamma}; \mu_0^{(c)} = \frac{\gamma - N}{\gamma} + (1 - \rho^2 + \beta); \\ \mu_i^{(m)} &= \mu_i^{(c)} = \frac{1}{2\gamma}, i = 1, \dots, 2N. \end{aligned} \quad (11)$$

The parameter γ controls the spread of the states around the mean, and it is calculated as $\gamma = \rho^2(N + \kappa)$, with $\kappa \geq 0$ to ensure semi-positive definiteness of the covariance matrix and

$0 \leq \rho \leq 1$. A good default choice is $\kappa = 0$ and ρ small enough to limit the spread of the sample states. The parameter β is introduced as a second control on the magnitude of the covariance weights. Details about the proper choice of κ , ρ and β can be found in van der Merwe (2004). The symbol $(\sqrt{\gamma \mathbf{P}})_i$ is the i th column (or row) of the root square matrix $\sqrt{\gamma \mathbf{P}}$, which can be regularly computed by Cholesky decomposition (e.g. Press et al., 1992).

The $2N+1$ sigma point vectors are assembled in the following matrix:

$$\mathcal{X}_{k-1} = \begin{bmatrix} \hat{\mathbf{x}}_{k-1} & \hat{\mathbf{x}}_{k-1} + \sqrt{\gamma \mathbf{P}_{k-1}} & \hat{\mathbf{x}}_{k-1} - \sqrt{\gamma \mathbf{P}_{k-1}} \end{bmatrix}. \quad (12)$$

As part of the prediction step, each sigma point vector is propagated through the dynamic state model:

$$\mathcal{X}_k^- = F_{k-1,k}(\mathcal{X}_{k-1}, \mathbf{u}_k). \quad (13)$$

The a priori estimate of the state mean is computed as the weighted average of the transformed points:

$$\hat{\mathbf{x}}_k^- = \sum_{i=0}^{2N} \mu_i^{(m)} \mathcal{X}_{k,i}^-. \quad (14)$$

The a priori estimate of the state covariance is computed as the weighted outer product of the transformed points plus the Gaussian noise covariance \mathbf{Q}_k :

$$\mathbf{P}_k^- = \sum_{i=0}^{2N} \mu_i^{(c)} \left(\mathcal{X}_{k,i}^- - \hat{\mathbf{x}}_k^- \right) \left(\mathcal{X}_{k,i}^- - \hat{\mathbf{x}}_k^- \right)^T + \mathbf{Q}_k. \quad (15)$$

One alternative to incorporate the effect of the process noise on the observed sigma-points, is to augment the number of these sigma-points with N additional vectors derived from the matrix square root of the process noise covariance, and recalculating the various weights μ_i accordingly. We opted to redraw the states, thus keeping the dimensionality of the problem equal to $2N+1$, although this option has the drawback that it discards the odd moments information captured by the propagated original sample states (van der Merwe, 2004):

$$\mathcal{X}_k^- = \begin{bmatrix} \hat{\mathbf{x}}_k^- & \hat{\mathbf{x}}_k^- + \sqrt{\gamma \mathbf{P}_k^-} & \hat{\mathbf{x}}_k^- - \sqrt{\gamma \mathbf{P}_k^-} \end{bmatrix}. \quad (16)$$

The observation equation is applied to this set of state vectors:

$$\mathcal{Y}_k = H_k(\mathcal{X}_k^-). \quad (17)$$

The forecast cross covariance between state predictions errors and observation predictions errors, $\mathbf{P}_{xy,k}$, and the forecast error covariance matrix of the observation predictions, $\mathbf{P}_{yy,k}$, are computed as follows:

$$\mathbf{y}_k^- = \sum_{i=0}^{2N} \mu_i^{(m)} \mathbf{y}_{k,i} \quad (18)$$

$$\mathbf{P}_{xy,k} = \sum_{i=0}^{2N} \mu_i^{(c)} (\mathbf{x}_{k,i} - \hat{\mathbf{x}}_k^-) (\mathbf{y}_{k,i} - \hat{\mathbf{y}}_k^-)^T \quad (19)$$

$$\mathbf{P}_{yy,k} = \sum_{i=0}^{2N} \mu_i^{(c)} (\mathbf{y}_{k,i} - \hat{\mathbf{y}}_k^-) (\mathbf{y}_{k,i} - \hat{\mathbf{y}}_k^-)^T \quad (20)$$

where $\hat{\mathbf{y}}_k^-$ is a weighted mean of the predicted measurements $\mathbf{Y}_{k,i}$.

The Kalman *gain* can be then straightforwardly computed with Eq. (4), while the a posteriori estimate of the state mean $\hat{\mathbf{x}}_k$ and the covariance matrix \mathbf{P}_k can be computed respectively with Eqs. (3) and (5).

In case of linear systems, the solution provided by UKF converges to that of SKF as the size of the sigma points ensemble increases (van der Merwe, 2004).

2.3 Ensemble Kalman Filter (EnKF)

The EnKF uses an ensemble of randomly chosen model trajectories, from which the necessary error covariances are estimated (Evensen, 2003). Similarly to the UKF, this method does not approximate the nonlinear process and observation models: it rather uses the true nonlinear models and it approximates the distribution of the state random variable. Moreover, it does not explicitly transform error information with a dynamic equation for computing the state error covariance matrix.

The EnKF propagates an ensemble of state vectors and each of these propagated vectors represents one realization of generated model replicas. Given an ensemble of L members, the dynamic model is applied to each member as follows:

$$\mathbf{x}_{k,i}^- = F_{k-1,k} (\mathbf{x}_{k-1,i}, \mathbf{u}_k) + \mathbf{v}_{k-1,i} \quad i = 1 \dots L \quad (21)$$

where $\mathbf{x}_{k,i}^-$ is the i th forecast ensemble member at time k and $\mathbf{x}_{k-1,i}$ is the updated ensemble member at $k-1$. Vector $\mathbf{v}_{k-1,i}$ is the i th column of a $N \times L$ matrix of perturbations generated according to a Gaussian distribution with zero mean and covariance \mathbf{Q}_k . In this study the current exogenous input vector is assumed to be unperturbed, to keep the analogy with the other two methods.

The sample mean and covariance can be evaluated according to the expressions:

$$\hat{\mathbf{x}}_k^- = \frac{1}{L} \sum_{i=1}^L \mathbf{x}_{k,i}^- \quad (22)$$

$$\mathbf{P}_k^- = \frac{1}{L-1} \sum_{i=1}^L (\mathbf{x}_{k,i}^- - \hat{\mathbf{x}}_k^-) (\mathbf{x}_{k,i}^- - \hat{\mathbf{x}}_k^-)^T. \quad (23)$$

In practice, the calculation of the approximate covariance \mathbf{P}_k^- is not required. The Kalman gain \mathbf{K} is obtained with Eq. (4) after computing the following covariances:

$$\mathbf{P}_{xy,k} = \frac{1}{L-1} \sum_{i=1}^L (\mathbf{x}_{k,i}^- - \hat{\mathbf{x}}_k^-) (\mathbf{y}_{k,i} - H_k (\hat{\mathbf{x}}_k^-))^T \quad (24)$$

$$\mathbf{P}_{yy,k} = \frac{1}{L-1} \sum_{i=1}^L (\mathbf{y}_{k,i} - H_k (\hat{\mathbf{x}}_k^-)) (\mathbf{y}_{k,i} - H_k (\hat{\mathbf{x}}_k^-))^T \quad (25)$$

where $\mathbf{y}_{k,i}^- = H_k (\mathbf{x}_{k,i}^-)$ represents the i th observation prediction at discrete time k .

An ensemble of L perturbed observations vectors $\mathbf{y}_{k,i}$ is derived by summing perturbations $\boldsymbol{\eta}_{k,i}$ ($\boldsymbol{\eta}_{k,i} \in N(0, \mathbf{R}_k)$, $i = 1 \dots L$) to the nominal term \mathbf{y}_k . The update step for the forecasted state ensemble members is then defined as follows:

$$\mathbf{x}_{k,i} = \mathbf{x}_{k,i}^- + \mathbf{K}_k (\mathbf{y}_{k,i} - H_k (\mathbf{x}_{k,i}^-)). \quad (26)$$

After the analysis ensemble is generated, it is propagated forward, and a new assimilation cycle starts.

The EnKF algorithm does not entail to explicitly update and store the state error covariance matrix, differently from the SKF and the UKF algorithms, where the state error covariance has to be explicitly updated to represent the change in forecast error covariance as an observation becomes available.

The ensemble forecast step of both UKF and EnKF algorithms can be parallelized by running each ensemble member on a separate processor of a parallel computer (or cluster). This can result in a significant computational advantage for the application of these two algorithms, whose computational effort is highly dependent on the size of the respective ensembles.

The optimal ensemble size is uncertain for the EnKF and it is generally heuristically chosen. Moreover, it is still not clear how the ensemble size should scale with the system dimension to achieve adequate estimates (Reichle et al., 2002a; Camporese et al., 2009). In case of linear systems, the solution of EnKF converges to that of SKF as the ensemble size increases (Evensen, 2003). For small ensemble size the EnKF is susceptible to systematic underestimation of the ensemble error covariance, caused by spurious long-range correlations (Houtekamer and Mitchell, 1998; Papadakis et al., 2010).

On the contrary, the UKF relies on a deterministically chosen set of samples to capture the statistical moments of the nonlinear model accurately, and the number of samples is univocally defined by the system dimension. However, the identification of this set of sample requires the calculation of the matrix square root of the state covariance matrix (see Eq. 11), which can be a computationally intensive process.

3 Soil water transport model

The soil water dynamics along the vertical direction is modelled using the Richards equation (Jury et al., 1991) in the h -based form:

$$C(h) \frac{\partial h}{\partial t} = \frac{\partial \left[K(h) \left(\frac{\partial h}{\partial z} + 1 \right) \right]}{\partial z} \quad (27)$$

where t is the time, z denotes the position along vertical axis (with upward orientation and zero reference value at the surface), h [L] is the matric pressure head [L], $K(h)$ [L T⁻¹] is the hydraulic conductivity function, and $C(h)$ [L⁻¹] is the differential water capacity function, obtained from the derivative $C(h) = d\theta(h)/dh$ of the water retention function $\theta(h)$ [L³ L⁻³].

The water retention and hydraulic conductivity functions are modelled according to the van Genuchten-Mualem model (van Genuchten, 1980):

$$\theta(h) = \theta_r + (\theta_s - \theta_r) \left[1 + |\alpha h|^n \right]^{-m} \quad (28)$$

$$K(\theta) = K_s \left(\frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^\lambda \left\{ 1 - \left[1 - \left(\frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^{1/m} \right]^m \right\}^2 \quad (29)$$

where θ_s [L³ L⁻³] is the saturated soil water content, θ_r [L³ L⁻³] is the residual soil water content, K_s [L T⁻¹] is the saturated hydraulic conductivity, while $\alpha > 0$ [L⁻¹], $n > 1$ [–], m [–] and λ [–] are empirical parameters. Following a common assumption, parameter m [–] is defined by the relation $m = 1 - 1/n$ and λ [–] is set equal to 0.5.

Equation (27), combined with the boundary conditions, can be solved by adopting a numerical scheme, which is formally equivalent to a state-space representation of the system model discretised in the time domain, as the one described by Eq. (1). Depending on the type of numerical scheme employed, the system model can be linear or nonlinear.

For this specific system, $\mathbf{v}(t)$ represents the state noise affecting the dynamic behaviour. According to Katul et al. (1993) the zero mean state noise assumption describes reasonably well the dynamic characteristics of the soil water flow in field conditions.

The measurement model could be reduced to a simple linear relation if matric pressure head is directly measured at given soil depths. If soil water content is measured, the observation equation is described by a nonlinear model corresponding to the soil water retention function (see Eq. 28). Different observation equations are required to assimilate other sources of measurements, such as those originated by near-surface remote sensing.

Below we illustrate three numerical schemes largely employed for integrating the Richards equation.

3.1 Explicit finite difference scheme (EX)

The Explicit finite difference scheme (EX) is the most basic numerical technique for solving differential equations. However, it may suffer from instability, which may make the method inappropriate or impractical. This method, also recognised as forward Euler finite difference scheme, is the one employed by Walker et al. (2001).

This numerical scheme provides an estimate of the matric head h_k^i of the i th node at the k th time-step as function of all other quantities at the preceding time-step according to the following discrete form of Eq. (27):

$$h_k^i = \left(\frac{\Delta t_{k-1}}{C_{k-1}^i} \frac{K_{k-1}^{i-1/2}}{\Delta z^i \Delta z^u}; 1 - \frac{\Delta t_{k-1}}{C_{k-1}^i} \frac{\frac{K_{k-1}^{i-1/2}}{\Delta z^u} + \frac{K_{k-1}^{i+1/2}}{\Delta z^l}}{\Delta z^i} \right. \\ \left. \frac{\Delta t_{k-1}}{C_{k-1}^i} \frac{K_{k-1}^{i+1/2}}{\Delta z^i \Delta z^l} \right) \begin{pmatrix} h_{k-1}^{i-1} \\ h_{k-1}^i \\ h_{k-1}^{i+1} \end{pmatrix} + \frac{\Delta t_{k-1}}{C_{k-1}^i} \frac{K_{k-1}^{i-1} - K_{k-1}^{i+1}}{2 \Delta z^i} \quad (30)$$

The subscript i for the node number is increasing downward. The soil column is divided in compartments of finite thickness Δz^i . All nodes, including the top and bottom nodes, are in the centre of the soil compartments, with $\Delta z^u = z^{i-1} - z^i$ and $\Delta z^l = z^i - z^{i+1}$. This represents a small difference with respect to the work of Walker et al. (2001), who considered nodes at the compartment extremes, positive upwards. $K_k^{i-1/2}$ and $K_k^{i+1/2}$ denote respectively the upward and the downward spatial averages of the hydraulic conductivity computed as arithmetic means. Δt_{k-1} indicates the time interval $\Delta t_{k-1} = t_k - t_{k-1}$.

For nodes $i = 1$ and $i = N$ the discrete forms at time-step k are:

$$h_k^1 = \left(\frac{\Delta t_{k-1}}{C_{k-1}^2} \frac{K_{k-1}^{1/2}}{\Delta z^2 \Delta z^l}; 1 - \frac{\Delta t_{k-1}}{C_{k-1}^2} \frac{\frac{K_{k-1}^{1/2}}{\Delta z^l} + \frac{K_{k-1}^{3/2}}{(z^2 - z^3)}}{\Delta z^2} \right. \\ \left. \frac{\Delta t_{k-1}}{C_{k-1}^2} \frac{K_{k-1}^{3/2}}{\Delta z^2 (z^2 - z^3)} \right) \begin{pmatrix} h_{k-1}^1 \\ h_{k-1}^2 \\ h_{k-1}^3 \end{pmatrix} + \frac{\Delta t_{k-1}}{C_{k-1}^2} \frac{K_{k-1}^1 - K_{k-1}^3}{2 \Delta z^2} - \Delta z^l \left(\frac{q_{\text{top}}}{K_{k-1}^{1/2}} + 1 \right) \quad (31)$$

$$h_k^N = \left(\frac{\Delta t_{k-1}}{C_{k-1}^{N-1}} \frac{K_{k-1}^{N-1-1/2}}{\Delta z^{N-1} (z^{N-2} - z^{N-1})}; 1 - \frac{\Delta t_{k-1}}{C_{k-1}^{N-1}} \right. \\ \left. \frac{\frac{K_{k-1}^{N-1-1/2}}{(z^{N-2} - z^{N-1})} + \frac{K_{k-1}^{N-1/2}}{\Delta z^u}}{\Delta z^{N-1}}; \frac{\Delta t_{k-1}}{C_{k-1}^{N-1}} \frac{K_{k-1}^{N-1/2}}{\Delta z^{N-1} \Delta z^u} \right) \\ \begin{pmatrix} h_{k-1}^{N-2} \\ h_{k-1}^{N-1} \\ h_{k-1}^N \end{pmatrix} + \frac{\Delta t_{k-1}}{C_{k-1}^{N-1}} \frac{K_{k-1}^{N-2} - K_{k-1}^N}{2 \Delta z^{N-1}} + \Delta z^u \left(\frac{q_{\text{bot}}}{K_{k-1}^{N-1/2}} + 1 \right). \quad (32)$$

The symbols q_{top} and q_{bot} indicate the top and bottom boundary conditions, herein assumed to be of Neumann type.

The forecasting equation of the system state \mathbf{x}_k , coinciding with the matric pressure head ($x_k^i = h_k^i$), can be obtained by combining the discrete equations written for all N nodes in the following linear state-space form:

$$\hat{\mathbf{x}}_k^- = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1} + \mathbf{g}_{k-1} \quad (33)$$

where \mathbf{A}_{k-1} is the matrix obtained by assembling the terms multiplying the state vector at time-step $k-1$ of Eqs. (30)–(32) and \mathbf{g}_{k-1} results from the combination of the terms on the right end of Eqs. (30)–(32).

According to this formulation, the system covariance is updated by the expression:

$$\mathbf{P}_k^- = \mathbf{A}_{k-1} \mathbf{P}_{k-1} \mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}. \quad (34)$$

Once the a priori state mean, $\hat{\mathbf{x}}_{k-1}^-$, and state covariance, \mathbf{P}_{k-1}^- , have been computed using Eqs. (33)–(34), the results of Eqs. (7)–(8) can be used for determining \mathbf{K} in Eq. (4), and then obtaining the a posteriori estimates $\hat{\mathbf{x}}_k$ and \mathbf{P}_k by means of Eqs. (3) and (5), respectively.

3.2 Crank-Nicolson finite difference scheme (CN)

The Crank-Nicolson implicit finite difference scheme (CN) has been widely implemented for solving the Richards equation (e.g. Haverkamp et al., 1977; Santini, 1980; Romano et al., 1998). The CN scheme is numerically stable, but sensitive to spurious oscillations when the ratio of the time step to the square of the space step is large.

The discrete representation of this relationship for intermediary nodes yields:

$$\begin{pmatrix} -\frac{K_{k-1}^{i-1/2}}{2 \Delta z^i \Delta z^u}; \frac{C_{k-1}^i}{\Delta t_{k-1}} + \frac{\frac{K_{k-1}^{i-1/2}}{\Delta z^u} + \frac{K_{k-1}^{i+1/2}}{\Delta z^l}}{2 \Delta z^i}; -\frac{K_{k-1}^{i+1/2}}{2 \Delta z^i \Delta z^l} \end{pmatrix} \begin{pmatrix} h_k^{i-1} \\ h_k^i \\ h_k^{i+1} \end{pmatrix} = \begin{pmatrix} \frac{K_{k-1}^{i-1/2}}{2 \Delta z^i \Delta z^u}; \frac{C_{k-1}^i}{\Delta t_{k-1}} - \frac{\frac{K_{k-1}^{i-1/2}}{\Delta z^u} + \frac{K_{k-1}^{i+1/2}}{\Delta z^l}}{2 \Delta z^i}; \frac{K_{k-1}^{i+1/2}}{2 \Delta z^i \Delta z^l} \end{pmatrix} \begin{pmatrix} h_{k-1}^{i-1} \\ h_{k-1}^i \\ h_{k-1}^{i+1} \end{pmatrix} + \frac{K_{k-1}^{i-1} - K_{k-1}^{i+1}}{2 \Delta z^i}. \quad (35)$$

For nodes $i=1$ and $i=N$ the discrete forms at time-step $k+1$ are:

$$\begin{aligned} & \begin{pmatrix} \frac{C_{k-1}^1}{\Delta t_{k-1}} + \frac{K_{k-1}^{\frac{1}{2}}}{2 \Delta z^1 \Delta z^l}; -\frac{K_{k-1}^{\frac{1}{2}}}{\Delta z^1 \Delta z^l} \end{pmatrix} \begin{pmatrix} h_k^1 \\ h_k^2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{C_{k-1}^1}{\Delta t_{k-1}} - \frac{K_{k-1}^{\frac{1}{2}}}{2 \Delta z^1 \Delta z^l}; \frac{K_{k-1}^{\frac{1}{2}}}{\Delta z^1 \Delta z^l} \end{pmatrix} \begin{pmatrix} h_{k-1}^1 \\ h_{k-1}^2 \end{pmatrix} + \frac{q_{\text{top}} - K_{k-1}^{\frac{1}{2}}}{\Delta z^1} \quad (36) \\ & \begin{pmatrix} -\frac{K_{k-1}^{N-1/2}}{\Delta z^N \Delta z^u}; \frac{C_{k-1}^N}{\Delta t_{k-1}} + \frac{K_{k-1}^{N-1/2}}{2 \Delta z^N \Delta z^u} \end{pmatrix} \begin{pmatrix} h_k^{N-1} \\ h_k^N \end{pmatrix} \\ &= \begin{pmatrix} \frac{K_{k-1}^{N-1/2}}{2 \Delta z^N \Delta z^u}; \frac{C_{k-1}^N}{\Delta t_{k-1}} - \frac{K_{k-1}^{N-1/2}}{2 \Delta z^N \Delta z^u} \end{pmatrix} \begin{pmatrix} h_{k-1}^{N-1} \\ h_{k-1}^N \end{pmatrix} + \frac{K_{k-1}^{-1/2} - q_{\text{bot}}}{\Delta z^N}. \quad (37) \end{aligned}$$

As in the previous algorithm, an explicit linearization of K and C is implemented, by taking their values at the previous

time-step $k-1$. Then, a linear state-space representation of the dynamic system can be easily derived by combining the set of Eqs. (35)–(37) written for each node and accounting for the boundary conditions:

$$\hat{\mathbf{x}}_k^- = (\mathbf{B}'_{k-1})^{-1} \mathbf{A}'_{k-1} \hat{\mathbf{x}}_{k-1} + (\mathbf{B}'_{k-1})^{-1} \mathbf{g}'_{k-1}. \quad (38)$$

This equation represents the dynamic state space model obtained with the CN scheme, analogously to Eq. (33) obtained with the explicit scheme. \mathbf{A}'_{k-1} and \mathbf{B}'_{k-1} are tri-diagonal matrices obtained by assembling the terms in the first parentheses on the right and left hand-sides of Eqs. (35)–(37), respectively. The term \mathbf{g}'_{k-1} is a vector obtained by assembling the terms on the right end of Eqs. (35)–(37).

The a priori estimate of the covariance matrix is calculated as follows:

$$\mathbf{P}_k^- = (\mathbf{B}'_{k-1})^{-1} \mathbf{A}'_{k-1} \mathbf{P}_{k-1} \left[(\mathbf{B}'_{k-1})^{-1} \mathbf{A}'_{k-1} \right]^T + \mathbf{Q}_{k-1}. \quad (39)$$

As for the explicit scheme, once the a priori system state and covariance have been determined, the current values $\hat{\mathbf{x}}_k$ and \mathbf{P}_k can be straightforwardly calculated with Eqs. (3)–(8).

3.3 Nonlinear implicit finite difference scheme (NL)

The backward Euler's finite differences (implicit) nonlinear (NL) scheme guarantees numerical stability for time steps considerably larger than those employed in the CN scheme, which explains its wide application in simulation models.

In this study we adopt the NL scheme of the Richards equation introduced by Celia et al. (1990) and further implemented by the SWAP model (van Dam, 2000).

The NL scheme has to be iteratively solved, as it includes C and θ values at the current time-step k to account for the strong nonlinearity of the differential water capacity C :

$$\begin{pmatrix} -\frac{K_{k-1}^{i-1/2}}{\Delta z^i \Delta z^u}; \frac{C_{k,p-1}^i}{\Delta t_{k-1}} + \frac{\frac{K_{k-1}^{i-1/2}}{\Delta z^u} + \frac{K_{k-1}^{i+1/2}}{\Delta z^l}}{\Delta z^i}; -\frac{K_{k-1}^{i+1/2}}{\Delta z^i \Delta z^l} \end{pmatrix} \begin{pmatrix} h_{k,p}^{i-1} \\ h_{k,p}^i \\ h_{k,p}^{i+1} \end{pmatrix} = \frac{C_{k,p-1}^i}{\Delta t_{k-1}} h_{k,p-1}^i + \frac{K_{k-1}^{i-1} - K_{k-1}^{i+1}}{2 \Delta z^i} + \theta_{k-1}^i - \theta_{k,p-1}^i. \quad (40)$$

The discrete forms for nodes $i=1$ and $i=N$ at time-step k are:

$$\begin{aligned} & \begin{pmatrix} \frac{C_{k,p-1}^1}{\Delta t_{k-1}} + \frac{K_{k-1}^{\frac{1}{2}}}{\Delta z^1 \Delta z^l}; -\frac{K_{k-1}^{\frac{1}{2}}}{\Delta z^1 \Delta z^l} \end{pmatrix} \begin{pmatrix} h_{k,p}^1 \\ h_{k,p}^2 \end{pmatrix} \\ &= \frac{C_{k,p-1}^1}{\Delta t_{k-1}} h_{k,p-1}^1 - \frac{q_{\text{top}} + K_{k-1}^{\frac{1}{2}}}{\Delta z^1} + \theta_{k-1}^1 - \theta_{k,p-1}^1 \quad (41) \end{aligned}$$

$$\begin{aligned} & \begin{pmatrix} -\frac{K_{k-1}^{N-1/2}}{\Delta z^N \Delta z^u}; \frac{C_{k,p-1}^N}{\Delta t_{k-1}} + \frac{K_{k-1}^{N-1/2}}{\Delta z^N \Delta z^u} \end{pmatrix} \begin{pmatrix} h_{k,p}^{N-1} \\ h_{k,p}^N \end{pmatrix} \\ &= \frac{C_{k,p-1}^N}{\Delta t_{k-1}} h_{k,p-1}^N + \frac{q_{\text{bot}} + K_{k-1}^{N-1/2}}{\Delta z^N} + \theta_{k-1}^N - \theta_{k,p-1}^N \quad (42) \end{aligned}$$

Table 1. Summary of assimilation algorithms adopted in combination with different numerical schemes and different types of observed variables.

Observed variable	Finite difference scheme			
	EX	CN	NL	NL
h	SKF ¹	SKF ¹	UKF ³	EnKF ⁴
θ	–	EKF ²	UKF ³	EnKF ⁴

¹ SKF = standard Kalman Filter; ² EKF = extended Kalman Filter; ³ UKF = unscented Kalman Filter; ⁴ EnKF = ensemble Kalman Filter.

where index p denotes the iteration step.

The dynamic state space model then assumes the form:

$$\mathbf{B}_{k-1,k}'' \hat{\mathbf{x}}_k^- = \mathbf{A}_{k-1,k}'' \hat{\mathbf{x}}_{k-1} + \mathbf{g}_{k-1,k}'' \quad (43)$$

$\mathbf{A}_{k-1,k}''$ is a diagonal matrix, with $\mathbf{A}_{k-1,k}^{i,i} = C_k^i / \Delta t_{k-1}$, obtained by assembling the terms multiplying the state vector at time-step $k-1$ in Eqs. (40)–(42). $\mathbf{g}_{k-1,k}''$ is a vector obtained by assembling the three terms on the right end of Eqs. (40)–(42). $\mathbf{B}_{k-1,k}''$ is the tri-diagonal matrix obtained by assembling the terms in the first parentheses on the left hand-side of Eqs. (40)–(42).

This type of numerical scheme requires the implementation of non-standard Kalman Filters for the a priori prediction of $\hat{\mathbf{x}}_{k-1}^-$ and \mathbf{P}_{k-1}^- . Equation (43) is employed for propagating an ensemble of sample states with both the UKF and the EnKF. In the case of the UKF, the a priori system state, $\hat{\mathbf{x}}_{k-1}^-$, and covariance, \mathbf{P}_{k-1}^- , are calculated as a weighted contribution of each propagated state, according to Eqs. (14)–(15). In the case of the EnKF, the statistics of the predicted states are computed with Eqs. (22)–(23).

4 Synthetic study

A synthetic study is performed to evaluate the relative merits of different Kalman Filter algorithms for retrieving matric head profiles by assimilating near surface soil matric head or water content measurements.

As pointed out above, the types of Kalman Filter that can be applied depend on the numerical scheme employed. The Standard Kalman Filter (SKF) can be implemented with explicit (EX) and Crank-Nicolson (CN) finite difference schemes, as long as the measurement model is linear. In this study, we employ the h -based form of the Richards equation and thus a linear measurement model occurs if the measured variable is also the matric pressure head h .

Walker et al. (2001) showed the efficiency of a standard Kalman Filter (SKF) in assimilating near surface matric pressure head measurements with an explicit finite difference scheme (EX) as compared with direct insertion of the observation values. Following Walker et al. (2001), we first

compare the SKF applied to a Crank-Nicolson finite difference scheme (SKF-CN), with the SKF applied to an explicit finite difference scheme (SKF-EX). Then we analyse the relative performances of the Unscented Kalman Filter based on the implicit nonlinear finite difference scheme (UKF-NL), the Ensemble Kalman Filter also implemented with the implicit numerical scheme (EnKF-NL) and the SKF-CN assimilation algorithm.

As an alternative to the UKF or the EnKF, the EKF could be also employed in conjunction with a nonlinear numerical scheme of the Richards equation. However, the EKF, based on an explicit linearization of nonlinear equations, is less efficient in state retrieving as compared with UKF (van der Merwe, 2004) and EnKF (e.g. Reichle et al., 2002b). Further discussions about the limitations and the flaws of the EKF can be found in other studies (e.g. Julier et al., 1995; van der Merwe, 2004).

If soil water content is the measured variable, the SKF-CN as such is not applicable, and a nonlinear KF is required to overcome the nonlinearity of the measurement model H_k defined by the soil water retention function. In this case we adopt an Extended approach (EKF-CN), using Eqs. (9)–(10) for computing the Kalman gain and the a posteriori estimate of the covariance matrix, respectively.

The final analysis of this work involves the comparison between the EKF-CN, the UKF-NL and the EnKF-NL algorithms, when the soil moisture content is the observation variable.

Table 1 summarizes the assimilation schemes examined in this paper.

The numerical experiment is arranged following Walker et al. (2001), to facilitate the comparison with this previous study. The essential information of the implemented numerical experiment is summarized in Table 2. Soil column depth is 100 cm, discretised in 27 nodes and the true initial matric pressure head profile is uniformly equal to -50 cm. The boundary conditions are: constant evaporative flux of $5.78 \times 10^{-6} \text{ cm s}^{-1}$ at the top surface and no flux at the bottom.

All assimilation scenarios are initialised with the same poor guess of the initial matric pressure head profile, assumed to be uniformly equal to -300 cm, thus 250 cm less than the true initial uniform profile.

Matric pressure head profiles are then retrieved by assimilating 4 different observation sets, consisting of hourly h data generated in top nodes, down to depths of 0.5, 1.5, 4.5 and 10 cm, respectively. These depths are slightly different from those adopted by Walker et al. (2001), because of the small differences in the soil column discretisation, as illustrated in Sect. 3.1. The NL scheme of the soil water transfer model illustrated above has been used to generate a set of soil water content and matric head profiles, representative of the true dynamic process to be retrieved. The assimilated values are true values of either matric heads or soil moisture contents,

Table 2. Parameters and conditions employed in the synthetic generation of matric pressure head profiles and for the initialization of the assimilation algorithms.

Soil depth	100 cm
Number of nodes	27
Soil hydraulic parameters	$\theta_s = 0.54 \text{ cm}^3 \text{ cm}^{-3}$ $\theta_r = 0.2 \text{ cm}^3 \text{ cm}^{-3}$ $\alpha = 0.008 \text{ cm}^{-1}$ $n = 1.8 (-)$ $K_s = 2.9 \times 10^{-4} \text{ cm}^{-1}$
Top evaporative flux	$5.79 \times 10^{-6} \text{ cm s}^{-1}$
Bottom flux	0 cm s^{-1}
Initial uniform h profile	-50 cm
Poor guess of initial uniform h profile	-300 cm
Initial state covariance matrix	$\mathbf{P}^{i,j} = \begin{cases} 10^3 \text{ cm}^2 & \text{if } i = j; i, j = 1 \dots n_{\text{nodes}} \\ 0 & \text{if } i \neq j \end{cases}$
Measurement noise variance matrix	$\mathbf{R}^{i,j} = \begin{cases} 0.02 y_i & \text{if } i = j; i, j = 1 \dots n_{\text{obs}} \\ 0 & \text{if } i \neq j \end{cases}$

perturbed with a random error of zero mean and a standard deviation of five percent of the absolute true values.

Employing a reference true dynamic scheme generated with a numerical scheme (NL in this case) different from that employed in the assimilation process (e.g. EX or CN) is formally equivalent to introducing a non-additive model error as further source of uncertainty in the system dynamics. However, we verified that the choice of the numerical scheme to be employed for generating the reference true dynamic process is not relevant for the comparison of the retrieved state profiles, since the profiles predicted with the three different numerical schemes are very similar to each other (herein not shown for the sake of brevity). Indeed, starting from the same initial conditions, the different numerical schemes (EX, CN and NL) can predict the evolution (a priori prediction) of the state profile with negligible numerical differences, as the time stepping adopted within each numerical scheme can be adjusted to guarantee numerical stability and achieve relatively similar numerical accuracy. This has been already proved by several previous studies, also applied for simulating infiltration fronts (e.g. Haverkamp et al., 1977; Paniconi et al., 1991; Kavetski et al., 2002), which from a numerical perspective is even more challenging than simulating an evaporation experiment, as the one examined in this study.

The EnKF has been implemented with an ensemble size of 50 members. We could verify that an ensemble size greater than 50 did not add much accuracy to the data assimilation algorithm, as also shown by Camporese et al. (2009) for the same case study. Moreover, by using 50 members for the examined case study, the dimensionality of the EnKF-NL assimilation algorithm is similar to that of the UKF-NL algorithm ($2N + 1 = 55$).

The default initial state variance is set equal to 10^3 cm^2 , rather than 10^6 cm^2 as assumed by Walker et al. (2001), since using an extremely high initial state variance causes practical

difficulties in the implementation of the UKF and the EnKF, as discussed later.

The amount of system noise variance is implemented in a different way. Walker et al. (2001) defined it equal to a five percent of the change in system states for each time step for the diagonal elements matrix, and zero for those off-diagonal. Entekhabi et al. (1994) considered an initial diagonal matrix accounting for the five percent of the precedent state. We adopted a diagonal system noise variance accounting for the five percent of the previous a posteriori state change, in order to avoid any ambiguity in the amount of error being incorporated, given that the time steps are variable in dependence of the chosen numerical scheme. The measurement noise covariance was set equal to two percent of the observations (either matric pressure head or soil moisture content) for the diagonal elements, and zero for all other elements, as also assumed by Walker et al. (2001).

4.1 Assimilating matric pressure heads

In Fig. 1, the profiles retrieved by assimilating daily matric pressure heads with the SKF-EX, SKF-CN, UKF-NL and EnKF-NL algorithms, are respectively compared with the true profiles as well as with the “guess” profiles. The “guess” profile, also referred to as “open loop” profile, is the one obtained without assimilating any near-surface observations, i.e. the system is simply propagated from the initial uniform conditions using the known boundary conditions.

All algorithms exhibit convergence rates of the retrieved matric head profiles to the true ones faster than that shown by Entekhabi et al. (1994), who implemented the EKF with a finite element scheme of the Richards equation for an analogous synthetic study.

The retrieved profiles tend to converge to the true ones slightly faster as the observation depth increases. For clarity, Fig. 1 illustrates only the results for the extreme observation

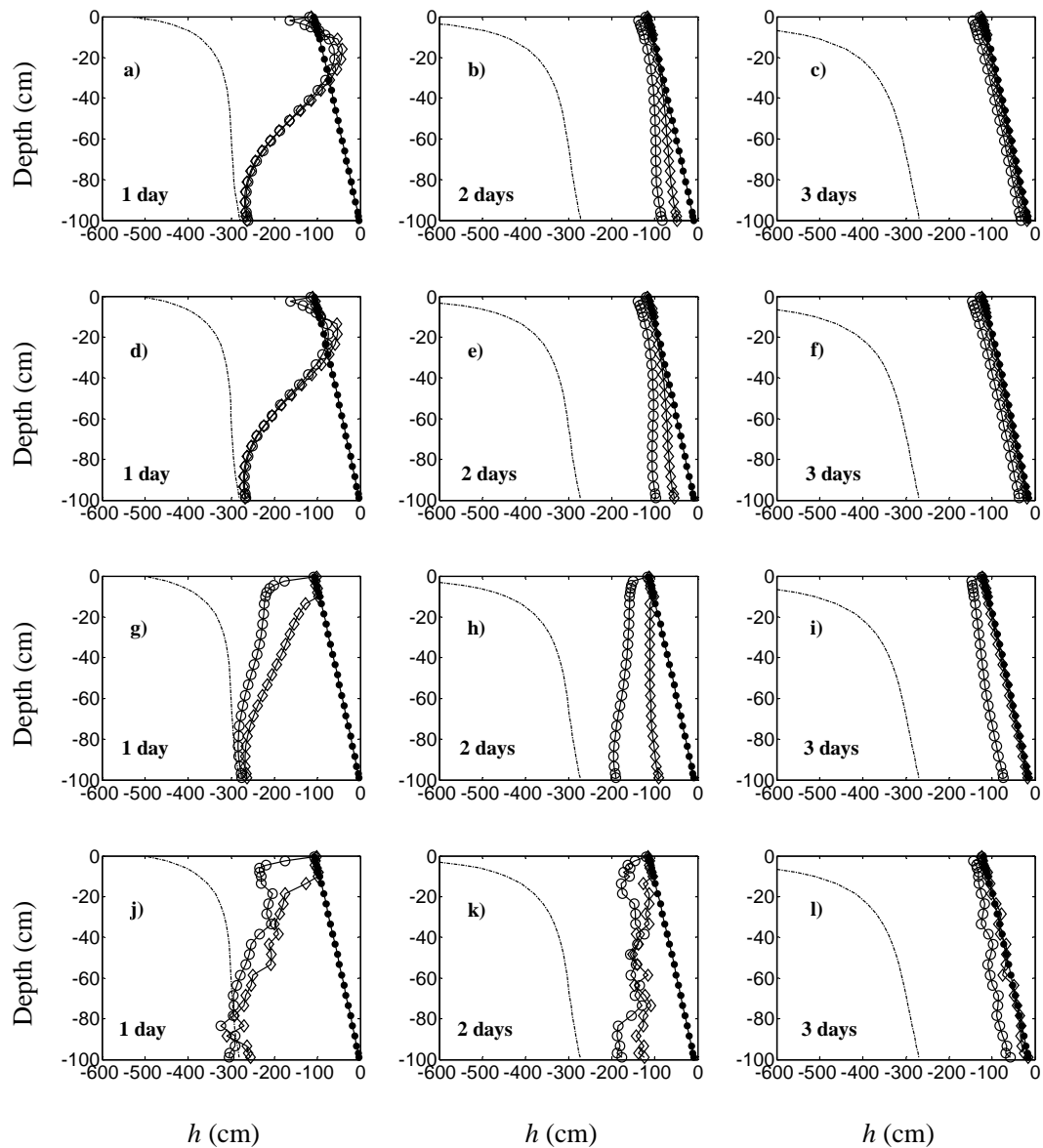


Figure 1. Retrieved profiles by assimilating daily observations of matric pressure heads involving nodes within the top 0.5 cm (open circle) and 10.5 cm (diamond) compared with the “true” profile (closed circle) and “guess” profile (dashed line), for different combinations of Kalman Filters and numerical schemes: (a–c) SKF-EX; (d–f) SKF-CN; (g–i) UKF-NL; (j–l) EnKF-NL.

depth scenarios considered in this work (i.e. observations at the nodes located within the top 0.5 and 10.5 cm, respectively), since the solutions for the other depth scenarios are bounded by those corresponding to these two extreme scenarios.

The two sets of profiles retrieved by using the SKF algorithm, respectively coupled with the explicit (Fig. 1a–c) and the Crank-Nicolson (Fig. 1d–f) numerical schemes, are practically identical. However, the explicit scheme requires time steps markedly smaller in order to guarantee numerical accuracy and stability. SKF-EX with a time step of 1 s demands about 30 times more CPU time than SKF-CN with a time step

of 60 s. The round off errors do not visibly impact the performance of the linearized approaches. Another favourable aspect of the CN numerical scheme is that it is particularly stable and thus it can be employed to assess the state profile even for a decreased precision of the model equation.

The profiles retrieved with UKF-NL and EnKF-NL are also very similar to each other, apart from the slight irregularities of the EnKF-NL profiles resulting from sample means (Eq. 22) obtained with a relatively small ensemble size. These irregularities tend to be smoothed as the ensemble size increases. Alternatively, they can be reduced by adopting specific sampling strategies both for the initial ensemble

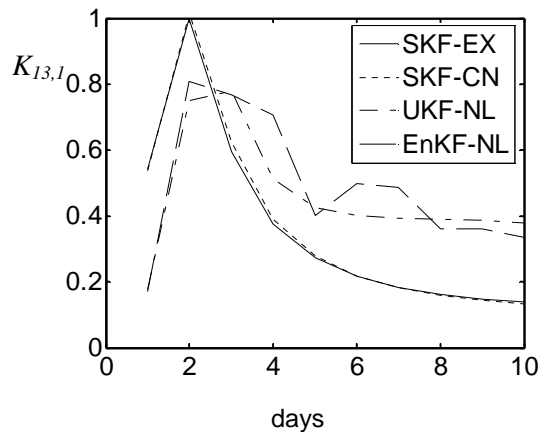


Figure 2. Evolution of the $K_{13,1}$ during the assimilation process with different assimilation algorithms.

and the measurement noise. These sampling strategies aim at generating ensembles with full rank and an improved conditioning with respect to what can be obtained with a pure random sampling, i.e. with a better representation of the error covariance matrix for a given ensemble size (Evensen, 2004).

The differences between the profiles retrieved with linear and nonlinear approaches are relatively high after the first update. UKF-NL and EnKF-NL first updates impact only slightly the medium and bottom nodes of the retrieved profile. Instead, using SKF-EX and SKF-CN a significant portion of the profile nodes is affected. These results are mainly related to differences in the prediction of the covariance \mathbf{P}^- prior to the update. The first order approximation of the EX and CN schemes induces a smoothing of the a priori covariance \mathbf{P}^- with respect to the NL scheme, which in turns also determines larger Kalman gain coefficients relating the state errors at middle depths to the observations at the top nodes, as illustrated in Fig. 2.

The generic Kalman gain coefficient $K_{i,j}$ is an index of the correlation between the state errors at the depth of the retrieved i th node and the observation at the depth of the assimilated j th node. Figure 2 shows the time variation of the $K_{13,1}$, i.e. the \mathbf{K} coefficient associated to the retrieval of the node at a depth of 48.5 cm (at about the middle of the soil column) by assimilating the top node observation. SKF-EX and SKF-CN exhibit patterns of $K_{13,1}$ different from those of UKF-NL and EnKF-NL, with values much larger on the first day of assimilation.

One of the key differences between EnKF or UKF and SKF is in the way the error covariance is propagated: while SKF propagates a single state vector (corresponding to the mean state) and, analytically, its error covariance, EnKF and UKF propagate an ensemble of state vectors, with error covariance resulting from the distribution of the states across the ensemble. However, the differences between the profiles

retrieved after the first update cannot be ascribed to the different strategies adopted by the examined algorithms in the propagation of the error covariance. EnKF and UKF, if applied to the EX and CN numerical schemes, would yield results close to those obtained with the SKF. Actually, in the limit of many simulations (i.e. by taking a large amount of members for the EnKF and a large amount of sigma points for the UKF), the corresponding EnKF and UKF estimates (state and error covariance) would tend to those provided by the SKF, which are statistically the optimal solutions for linear systems, provided that no model errors occur, the entering noise is white and the noise covariances are known.

Moreover, the differences between the profiles retrieved after the first update cannot be ascribed to discrepancies between the a priori state profile $\hat{\mathbf{x}}_k^-$ predicted by the different numerical schemes, since these discrepancies are negligible. Starting from the same initial poor guess, linear (EX, CN) and nonlinear numerical (NL) schemes predict the evolution of the state profile (a priori prediction) with very small differences.

In the UKF-NL and EnKF-NL algorithms, the nonlinear implicit differential scheme of the Richards equation is solved for each sample to predict its state evolution, while the update phase is activated only as the observation is available. In this numerical experiment, the NL differential scheme is resolved with an hourly time-step, thanks to its high numerical stability and accuracy, demanding a CPU time equal to almost 50 % that required for the CN scheme, which is implemented with a time step of 200 s. However, the nonlinear data assimilation approaches (UKF-NL and EnKF-NL) consume almost 50 % more CPU time than the SKF-CN, because of the cost of the ensemble propagation. The EnKF-NL is just slightly more efficient than the UKF-NL in terms of computational costs.

The effects of the system noise variance and of the initial state covariance deserve some comments.

Walker (1999) found that the performance of the retrieval algorithm is not particularly sensitive to the system noise variance. However, this result should be interpreted keeping in mind that Walker (1999) assumed the system noise variance to be equal to five percent of the state change. Instead, in the present study, the system noise variance has a more important role, as it is assumed to be equal to five percent of the previous state, similarly to Entekhabi et al. (1994).

As also shown by Walker (1999) in a sensitivity analysis for the same experimental setup, the convergence time tends to decrease as one takes higher initial state covariance values. With a value of 10^4 cm^2 applied to the diagonal elements of the initial state covariance, the retrieved profiles perfectly coincide with the true ones on third day of assimilation, regardless of the observation depth. Instead, as shown by Fig. 1, with a value of 10^3 cm^2 applied to the diagonal elements of the initial state covariance, perfect coincidence is achieved on the third day of assimilation only with an observation depth of 10.5 cm.

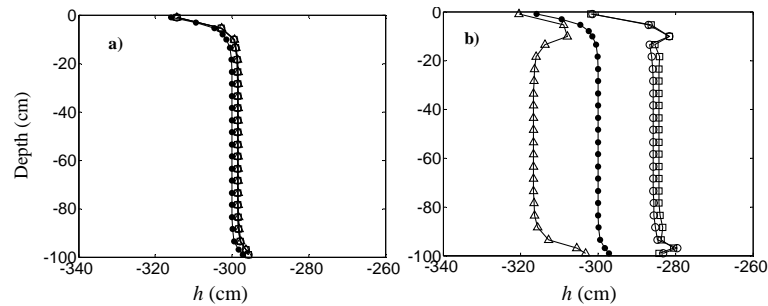


Figure 3. UKF-NL a priori estimates of the state mean after 1 h, computed with $\rho = 0.05$ (open circle), $\rho = 0.3$ (open square) and $\rho = 0.8$ (triangle), and adopting initial variances respectively equal to (a) 10^3 cm^2 and (b) 10^4 cm^2 on the diagonal elements of the state covariance matrix. The figure also shows the corresponding central state profiles (closed circle) propagated through the dynamic model on the same simulation time.

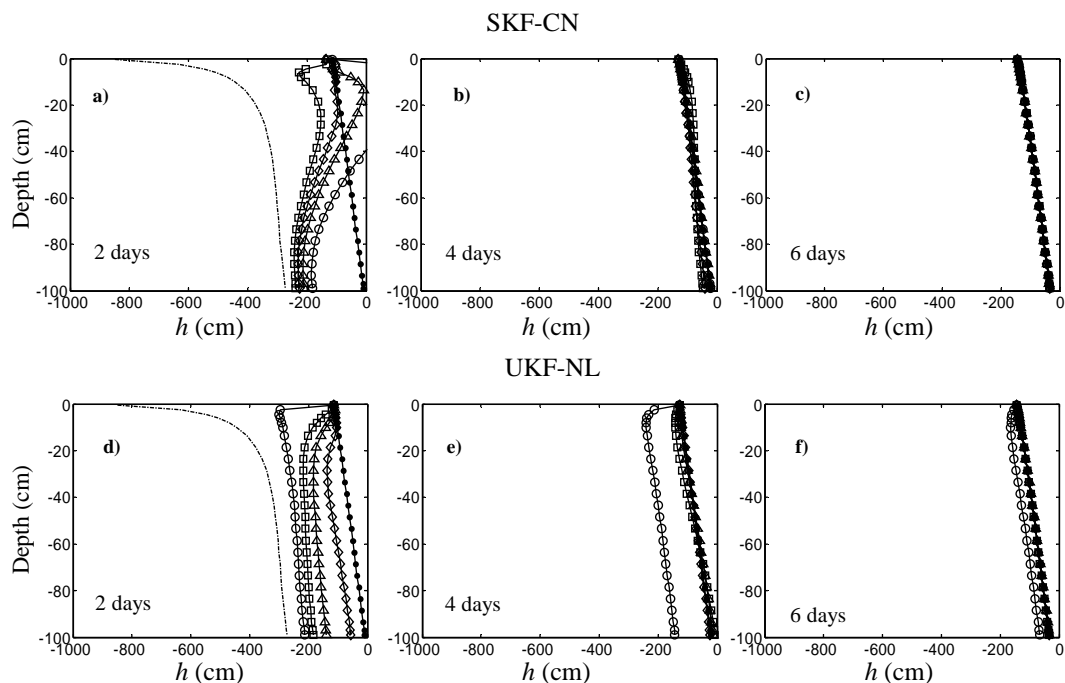


Figure 4. Retrieved profiles after (a, d) 2 days; (b, e) 4 days, and (c, f) 6 days using SKF-CN and UKF-NL, respectively, by assimilating matric pressure head observations every two days involving the nodes within the top 0.5 cm (open circle), 1.5 cm (square), 4.5 cm (triangle) and 10.5 cm (diamond), as compared with the “true” profile (closed circle) and “guess” profile (dashed line).

The choice of high initial state variance can actually determine some difficulties in the implementation of the UKF and EnKF algorithms. In both these algorithms, the distribution of the sample states logically depends on the magnitude of the variance. Taking a very large initial variance, without any correlation structure, could lead to sample profiles which are physically improbable, and consequently the stability of the assimilation process could degenerate. In case of the UKF, this issue can be overcome by shrinking the sample state distribution around the mean state with the scaling parameter ρ , which controls the weights attributed to the sample state distribution (Eq. 11). Figure 3 shows different

UKF-NL a priori estimates of the state mean \hat{x}_k^- , computed after 1 h with Eq. (14), by adopting parameter ρ respectively equal to 0.05, 0.3 and 0.8, and initial variances respectively equal to 10^3 and 10^4 cm^2 for the diagonal elements of the state covariance matrix. As a reference state profile for the comparative analysis of the different a priori estimates, Fig. 3 shows the corresponding central state profile (X_0 of Eq. 11) propagated through the dynamic model (Eq. 13) on the same simulation time. The a priori mean state estimation using the smaller initial state variance value is practically insensitive to the value of ρ , while this prior mean is highly affected by ρ when using the higher initial state variance. Taking $\rho = 0.8$,

with a uniform initial profile -300 cm and state variance of 10^4 cm², leads to a set of sample profiles exhibiting positive matric heads for perturbed nodes. All positive matric heads were changed to zero before propagating the state through the dynamic model, thus favouring a shift of the predicted prior mean state towards smaller matric heads as compared with the central state. With ρ equal to 0.05 and 0.3 the saturated region is not sampled, but the predicted prior means are shifted towards larger matric heads with respect to the central state.

In practical applications, particularly those where the frequency of the observations is small, the evolution of the magnitudes of the state variance is unpredictable. Thus, the coefficients selected for designing the sigma point sampling strategy could turn out to be inadequate during the assimilation process. To overcome this issue, some other UKF applications adapt the value of the coefficients to guarantee the physical coherence of the sample states.

Figure 4 shows the profiles retrieved by assimilating matric head observations once every two days with the SKF-CN and the UKF-NL. In both cases the initial state covariance employed has been set equal to 10^3 cm². To guarantee a performance of the UKF-NL comparable with that of SKF-CN, the parameter ρ has been specifically tuned. Results show that the SKF-CN is able to retrieve the true profile already at the fourth day, i.e. at the second assimilation, practically for any observation depths. However, the profile retrieved with an observation depth of 0.5 cm exhibits positive matric heads in the upper 40 cm on the second day of assimilation. Satisfactory results with the SKF have also been obtained (not shown here) by assimilating observations once every three days and four days, except for the case when the observations are limited to the top node. SKF failed only for larger observation time-intervals, as the extremely negative pressure heads of the predicted states at the top nodes altered the singularity of the matrix operators. To achieve results similar to those obtained with SKF, specific tunings of the assimilation scheme are required with the UKF.

However, an aspect favouring the application of the UKF or the EnKF is the possibility to straightforwardly implement nonlinear state variable transformations. When the states are very far from the observations, the filtering process imposes severe gradients in the profiles, causing the estimation of temporary meaningless profiles. This could be partially avoided by making a transformation of the state (e.g. logarithmic), which implies at the same time a nonlinear transformation of the dynamic equation. This transformation, scarcely affecting results (not shown) in case of UKF or EnKF, should be treated by linearizing an already linearized equation in the case of an Extended Kalman Filter, therefore with a drastic reduction of the efficiency of the assimilation algorithm.

4.2 Assimilating of soil water content observations

The analysis has also been extended to examine the case of assimilating soil water content observations, instead of matric pressure heads.

The CN numerical scheme has been coupled with the EKF in order to handle the nonlinearity of the soil water retention function. Thus, the profiles retrieved with the EKF-CN algorithm have been compared with those retrieved with the UKF-NL and the EnKF-NL algorithms. An approach alternative to EKF-CN could have been to combine the SKF-CN algorithm for estimating the a priori mean state and covariance matrix of the state distribution, and apply the UKF just for the update phase when the soil moisture content observation becomes available, in order to address the nonlinearity of the water retention curve.

The EKF-CN algorithm entails the Jacobian matrices of the observation equation for computing the Kalman gain and the a posteriori error covariance, according to Eqs. (9)–(10). In this case study, matrix \mathbf{D}_k is the identity matrix, while the non-zero coefficients of \mathbf{C}_k coincide with the specific water capacity values computed at the of the observation nodes with respect to the a priori estimate $\hat{\mathbf{x}}_k^-$. Provided that the VGM model equations are continuous over the entire range of matric heads, the specific water capacity can be analytically derived. Thus the computation of the Kalman gain (Eq. 9) and of the a posteriori error covariance (Eq. 10) is very easy to be implemented and does not represent a significant increase of the overall computational load. Ultimately, the computational effort in the implementation of EKF-CN is not significantly different from that of SKF-CN for the examined system.

Figure 5 is designed to compare the matric head profiles retrieved with the EKF-CN, UKF-NL and EnKF-NL, analogously to Fig. 1, but assimilating soil water content observations instead of matric heads. The performances of the EnKF-NL and the UKF-NL algorithms are very similar to each other in terms of accuracy of the retrieved profiles, analogously to what observed with the same algorithms applied for assimilating matric pressure heads. The EnKF-NL is just slightly less onerous in terms of computational costs. Moreover, the profiles retrieved with the EKF-CN algorithm are markedly different from those obtained with the NL-based ones, particularly on the first day of assimilation. The EKF-CN algorithm requires a CPU time about twice smaller than the EnKF-NL and UKF-NL algorithms, because more time is required to ensure the convergence of the iterative numerical schemes implemented in the EnKF-NL and UKF-NL algorithms for each ensemble member.

The effect of the observation depth is more evident when taking soil water content instead of matric head as observation variable. For an observation depth of 10.5 cm (diamonds in Figs. 1 and 5), the convergence rate of the retrieved profiles towards the true ones by assimilating soil water content is very similar to that obtained by assimilating pressure

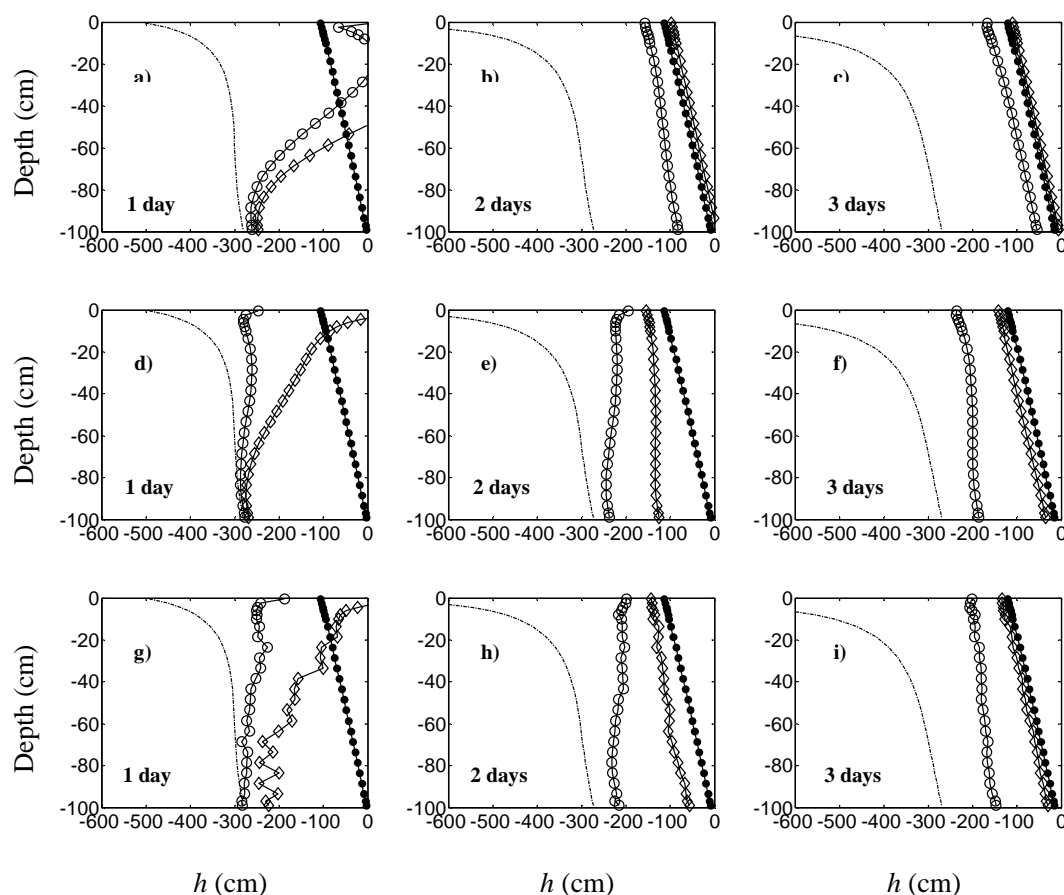


Figure 5. Retrieved profiles by assimilating daily observations of soil moisture contents involving nodes within the top 0.5 cm (open circle) and 10.5 cm (diamond) compared with the “true” profile (closed circle) and “guess” profile (dashed line), for different combinations of Kalman Filters and numerical schemes: (a–c) EKF-CN; (d–f) UKF-NL; (g–i) EnKF-NL.

heads. The profiles retrieved with an observation depth of 0.5 cm (open circles in Fig. 5) converge towards the true ones much more slowly than those retrieved with an observation depth of 10.5 cm, particularly for the UKF-NL and EnKF-NL schemes, for which the convergence is far from being reached even on the third day.

The differences between the matric head profiles retrieved by assimilating soil water contents instead of matric heads are entirely related to the nonlinearity of the observation equation, represented by the water retention function, which produces a deformation of the Kalman gain matrix with respect to that derived by assimilating matric heads, influencing both the state and the error covariance updates. This effect can be more easily interpreted by comparing the state profiles respectively retrieved with EKF and SKF, which show differences that are directly related to the linearization of the observation equation (first order approximation) by means of the specific water capacity.

The effect of this first order approximation is expected to be dependent on the soil properties. Coarser textured soils exhibit large gradients in the specific water capacity functions

for near-saturated conditions and thus are more vulnerable to biased predictions of the linearized observation equation. This effect can be exacerbated in dynamic processes involving infiltration, which impose limited correlation between nodes located at the opposite ends of the infiltration front, particularly in applications with real data, where model errors due to the uncertainty in the soil parameters are unavoidable.

It is important to point out that the results obtained by assimilating soil water content values, are subjected to the assumption that the parameters defining the soil water retention at the observation points coincide with those employed for simulating the soil water dynamics along the entire soil column, which is considered homogenous. However, in realistic circumstances, the retrieval algorithm should account for the model simplifications and system heterogeneity, particularly for the large spatial variability of the soil hydraulic properties (Pringle et al., 2007; Chirico et al., 2010). The “optimal” parameters defining the soil water retention at the observation points are in principle different from the “optimal” parameters defining the soil water dynamics along the soil column.

These “optimal” soil hydraulic parameters are to be considered as effective values at the scale of the observation or of the modelled system, respectively (Vereecken et al., 2007).

5 Conclusions

When designing a Kalman Filter algorithm for assimilating near surface data into the Richards equation, it would be desirable to choose the Kalman Filter considering the numerical scheme employed for solving the Richards equation and the type of assimilated variable.

It is well known that the nonlinear (NL) numerical scheme is computationally more efficient than the linearised Crank–Nicolson (CN) numerical scheme and for this reason the NL scheme is preferred by the most popular soil water dynamics simulation models. However, in some circumstances the larger computational effort required for the application of linear schemes is compensated by applying SKF analytic estimators for the propagation of the first two moments of the state profiles, with minor computational costs than the corresponding non-standard KF algorithms, without losing accuracy of the predictions.

In this study we verified this possibility by assimilating surface matric head or soil water content observations for retrieving matric head profiles of a homogeneous soil column subjected to a surface evaporation flux.

When assimilating matric heads, the standard Kalman Filter (SKF) combined with a Crank–Nicolson (CN) linear numerical scheme of the h -based form of the Richards equation, provides estimates of the matric head profiles converging faster to the true solution and with minor computational costs than the ensemble Kalman Filter (EnKF) or the unscented Kalman Filter (UKF) with a NL scheme.

The SKF-CN algorithm takes advantage of both the stability of the CN numerical scheme and the linearity of the operators in the dynamic system model. An explicit numerical scheme, although being also linear, is unfeasible for practical applications, as it demands computational time-steps of the order of a few seconds.

The SKF-CN becomes unfeasible for large scale applications, due to the computational issues attached to an explicit formulation of the covariance propagation using the SKF (Reichle et al., 2002a; Reichle, 2008; Camporese et al., 2009). However, even large scale applications can often be reduced to a set of independent low-dimensional spatial systems. Reichle and Koster (2003) compared the results of an Ensemble Kalman Filter (EnKF) applied to a 1-D model to those obtained with EnKF applied to a 3-D model to estimate soil moisture in the root zone, finding that, although the 3-D approach produced more accurate results, in particular for intermediate assimilated data volumes, the 1-D EnKF satisfactorily performed with a lower demand of computational time. Using either a low volume of assimilated data or a high number of them, the differences were not very pronounced.

The relative gain of using the 1-D EnKF over the straight model (with no assimilation) can be larger than the relative gain of using the 3-D EnKF over the 1-D EnKF, at least for those circumstances where horizontal error correlations have a marginal role (Hoebein and Troch, 2000; Reichle and Koster, 2003).

In order to make the SKF-CN applicable, the form of the Richards equation should be chosen according to the assimilated variable, so that the type of variable describing the observations is equal to that describing the states or it is at least a linear transformation of it. However, this strategy is not always possible, as for example it occurs when the assimilation algorithm has to be implemented with closed on-hand model software, such as HYDRUS (Vogel et al., 1996), with a pre-defined nonlinear numerical scheme.

Another case is when the observed variable is soil water content while the Richards equation has to be in the h -based form in order to handle both saturated and unsaturated flows in the vadose zone, thus imposing a nonlinear observation equation in the state-space description of the dynamic system. In this case, the SKF can be replaced with the Extended Kalman Filter to handle the nonlinearity of the observation equation. With regards to the numerical evaporation test examined in this study, we could verify that, when assimilating surface soil moistures for retrieving matric head profiles, the Extended Kalman Filter combined with the Crank–Nicolson (EKF-CN) also provides state profiles converging faster to the true profiles than the UKF-NL or EnKF-NL numerical schemes, with less computational efforts. However, we also argued that the linearization of the observation equation embedded in EKF does not ensure accurate solutions for all circumstances, particularly for coarser textured soils and when sharp space-time gradients of the state vector are involved in the retrieving process.

We also compared the EnKF with the UKF, both as potential alternatives to the EKF, whose flaws and limitations for nonlinear models have been largely discussed in the literature (e.g. Julier et al., 1995; van der Merwe, 2004). The UKF has been less commonly applied in hydrological studies and, to our knowledge, previous studies illustrating the implementation of the UKF with a numerical scheme of the Richards equation have not been published.

Differently from the SKF, both UKF and EnKF sampling strategies can lead to state profiles with low physical meaning, particularly with large state covariances.

The computational efforts for implementing EnKF-NL and UKF-NL are directly linked to the corresponding sizes of the state ensemble propagated during the retrieving process. For the UKF, the size of the propagated ensemble is deterministically set equal to $2N + 1$, where N is the system dimension. The optimal ensemble size for EnKF is not known a priori and it has been often calibrated, making it dependent on both the type of examined process and the system dimension (Reichle et al., 2002a). However, UKF also requires the calibration of a scaling parameter affecting the spread of the

sample state distribution around the central state, which can play a critical role in the overall retrieving process.

At least for the numerical evaporation test examined in this study, as far as the system dimension is almost half the ensemble size involved in EnKF, the two algorithms are almost equivalent in terms of computational costs. The UKF can be even more competitive than the EnKF, if the dimensionality on the nonlinear problem is much smaller than 50, which is the ensemble size generally required to gain accurate solutions in this type of applications (Camporese et al., 2009).

While dealing with different types of filters applied to nonlinear systems, it is important to keep in mind that the physical model affects the ensemble generation and the uncertainty analysis, which can be relevant for skewed ensemble distributions (Drécourt, 2004), as observed by Reichle et al. (2002a) for very dry or wet soil conditions. Ultimately, as these Authors uphold, the “best” approach for a given data assimilation problem will be application dependent.

This study focussed on retrieving state profiles, while assuming that parameters are known. However, in most practical circumstances, significant uncertainties arise from the identification of the soil hydraulic parameters. The following two companion papers (Medina et al., 2013a, b) explore the capability of a dual Kalman Filter approach for simultaneous retrieval of states and parameters in the Richards equation, by examining synthetic and experimental data, respectively.

Acknowledgements. This study has been supported by P.O.N. project “AQUATEC – New technologies of control, treatment, and maintenance for the solution of water emergency”. Hanoi Medina has been also supported by the Abdus Salam International Centre for Theoretical Physics (ICTP), where he has been appointed as Junior Associate. The Editor and the anonymous reviewers contributed considerably to improve the manuscript with fruitful comments.

Edited by: J. Vrugt

References

- Camporese, M., Paniconi, C., Putti, M., and Salandin, P.: Ensemble Kalman filter data assimilation for a process based catchment scale model of surface and subsurface flow, *Water Resour. Res.*, 45, W10421, doi:10.1029/2008WR007031, 2009.
- Celia, M. A., Bououtas, E. T., and Zarba, R. L.: A general mass-conservative numerical solution for the unsaturated flow equation, *Water Resour. Res.*, 26, 1483–1496, 1990.
- Chirico, G. B., Medina, H., and Romano, N.: Functional evaluation of PTF prediction uncertainty: An application at hillslope scale, *Geoderma*, 155, 193–202, 2010.
- Clark, M., Rupp, D., Woods, R., Zheng, X., Ibbitt, R., Slater, A., Schmidt, J., and Uddstrom, M.: Hydrological data assimilation with the ensemble Kalman filter: use of streamflow observations to update states in a distributed hydrological model, *Adv. Water Resour.*, 31, 1309–1324, 2008.
- Das, N. N. and Mohanty, B. P.: Root zone soil moisture assessment using remote sensing and vadose zone modeling, *Vadose Zone J.*, 5, 296–307, 2006.
- De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, *Water Resour. Res.*, 43, W09410, doi:10.1029/2006WR005449, 2007.
- Drécourt, J. P.: Data assimilation in hydrological modeling, Ph. D. Thesis, Environment & Resources DTU, Technical University of Denmark, Kongens Lyngby, Denmark, 2004.
- Dunne, S. and Entekhabi, D.: Land surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment, *Water Resour. Res.*, W01407, doi:10.1029/2005WR004334, 2006.
- Entekhabi, D., Nakamura, H., and Njoku, E. G.: Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remote sensed observations, *IEEE T. Geosci. Remote*, 32, 438–448, 1994.
- Entekhabi, D., Rodríguez-Iturbe, I., and Castelli, F.: Mutual interaction of soil moisture state and atmospheric processes, *J. Hydrol.*, 184, 3–17, 1996.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 10143–10162, 1994.
- Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynam.*, 53, 343–367, 2003.
- Evensen, G.: Sampling strategies and square root analysis schemes for the EnKF, *Ocean Dynam.*, 54, 539–560, 2004.
- Haverkamp, R. M., Vauclin, M., Tournia, J., Wierenga, P. J., and Vachaud, G.: A comparison of numerical simulation models for one-dimensional infiltration, *Soil Sci. Soc. Am. J.*, 41, 285–294, 1977.
- Heathman, G. C., Starks, P. J., Ahuja, L. R., and Jackson, T. J.: Assimilation of surface soil moisture to estimate profile soil water content, *J. Hydrol.*, 279, 1–17, 2003.
- Hoeben, R. and Troch P. A.: Assimilation of active microwave observation data for soil moisture profile estimation, *Water Resour. Res.*, 36, 2805–2819, doi:10.1029/2000WR900100, 2000.
- Houtekamer, P. L. and Mitchell, H.: Data assimilation using an ensemble Kalman filter technique, *Mon. Weather Rev.*, 126, 796–811, 1998.
- Huang, C., Li, X., Lu, L., and Gu, J.: Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter, *Remote Sens. Environ.*, 112, 888–900, doi:10.1016/j.rse.2007.06.026, 2007.
- Julier, S. J. and Uhlmann, J. K.: A New Extension of the Kalman Filter to Nonlinear Systems, *Proc. SPIE*, 3068, 182–193, 1997.
- Julier, S. J. and Uhlmann, J. K.: Unscented filtering and nonlinear estimation, *Proc. IEEE*, 92, 401–422, 2004.
- Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F.: A new approach for filtering nonlinear systems, in: *The Proceedings of the American Control Conference*, Seattle WA, USA, 1628–1632, 1995.
- Jury, W. A., Gardner, W. R., and Gardner, W. H.: *Soil Physics*, 5th Edn., John Wiley, New York, 1991.
- Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, *ASME J. Basic Eng.*, 82D, 35–45, 1960.

- Katul, G. G., Wendroth, O., Parlange, M. B., Puente, C. E., Folegatti, M. V., and Nielsen, D. R.: Estimation of in situ hydraulic conductivity function from nonlinear filtering theory, *Water Resour. Res.*, 29, 1063–1070, 1993.
- Kavetski, D., Binning, P., and Sloan, S. W.: Noniterative time stepping schemes with adaptive truncation error control for the solution of Richards equation, *Water Resour. Res.*, 38, 1211, doi:10.1029/2001WR000720, 2002.
- Luo, X. and Moroz, I. M.: Ensemble Kalman filter with the unscented transform, *Physica D*, 238, 549–562, 2009.
- Luo, X. and Moroz, I. M.: Reply to “Comment on ‘Ensemble Kalman filter with the unscented transform’”, *Physica D*, 239, 1662–1664, 2010.
- Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V. R. N., De Lannoy, G. J. M., De Keyser, R., and Savenije, H. H. G.: Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept, *Hydrol. Earth Syst. Sci.*, 14, 1773–1785, doi:10.5194/hess-14-1773-2010, 2010.
- McLaughlin, D. B.: An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering, *Adv. Water Resour.*, 25, 1275–1286, 2002.
- Medina, H., Romano, N., and Chirico, G. B.: Kalman filters for assimilating near-surface observations into the Richards equation – Part 2: A dual filter approach for simultaneous retrieval of states and parameters, *Hydrol. Earth Syst. Sci.*, 18, 2521–2541, doi:10.5194/hess-18-2521-2014, 2014.
- Medina, H., Romano, N., and Chirico, G. B.: Kalman filters for assimilating near-surface observations into the Richards equation – Part 3: Retrieving states and parameters from laboratory evaporation experiments, *Hydrol. Earth Syst. Sci.*, 18, 2543–2557, doi:10.5194/hess-18-2543-2014, 2014.
- Milly, P. C. and Dunne, K. A.: Sensitivity of the global water cycle to the water-holding capacity of land, *J. Climate*, 7, 506–526, 1994.
- Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P.: Dual state-parameter estimation of hydrological models using Ensemble Kalman filter, *Adv. Water Resour.*, 28, 135–147, 2005.
- Paniconi, C., Aldama, A. A., and Wood, E. F.: Numerical evaluation of iterative and noniterative methods for the solution of the nonlinear Richards equation, *Water Resour. Res.*, 27, 1147–1163, doi:10.1029/91WR00334, 1991.
- Papadakis, N., Mémin, E., Cuzol, A., and Gengembre, N.: Data assimilation with the weighted ensemble Kalman filter, *Tellus A*, 62, 673–697, 2010.
- Pipunic, R. C., Walker, J. P., and Western, A.: Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study, *Remote Sens. Environ.*, 112, 1295–1305, 2008.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Edn., Cambridge University Press, 1992.
- Pringle, M. J., Romano, N., Minasny, B., Chirico, G. B., and Lark, R. M.: Spatial evaluation of pedotransfer functions using wavelet analysis, *J. Hydrol.*, 333, 182–198, 2007.
- Reichle, R. H.: Data assimilation methods in the Earth sciences, *Adv. Water Resour.*, 31, 1411–1418, 2008.
- Reichle, R. H. and Koster, R. D.: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation, *J. Hydrometeorol.*, 4, 1229–1242, 2003.
- Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, *Mon. Weather Rev.*, 130, 103–114, 2002a.
- Reichle, R. H., Walker, J. P., Koster, R. D., and Houser, P. R.: Extended versus ensemble Kalman filtering for land data assimilation, *J. Hydrometeorol.*, 3, 728–740, 2002b.
- Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *J. Geophys. Res.-Atmos.*, 112, D09108, doi:10.1029/2006JD008033, 2007.
- Rodriguez-Iturbe, I. and Porporato, A.: *Ecohydrology of Water-Controlled Ecosystems: Soil Moisture and Plant Dynamics*, Cambridge University Press, 2005.
- Romano, N., Brunone, B., and Santini, A.: Numerical analysis of one-dimensional unsaturated flow in layered soils, *Adv. Water Resour.*, 21, 315–324, 1998.
- Santini, A.: Model for simulating soil water dynamics considering root extraction, *Proceedings of CCE Seminaires sur l'Irrigation Localisee – Agrimed Sorrento*, Italy, 1980.
- van Dam, J. C.: Field-scale water flow and solute transport. SWAP model concepts, parameter estimation and case studies, *Doctoral Thesis Wageningen University*, Wageningen, 2000.
- van der Merwe, R.: *Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models*, Ph. D. dissertation, University of Washington, USA, 2004.
- van Genuchten, M. Th.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, 44, 892–898, 1980.
- Vereecken, H., Kasteel, R., Vanderborght, J., and Harter T.: Upscaling hydraulic properties and soil water flow processes in heterogeneous soils, *Vadose Zone J.*, 6, 1–28, doi:10.2136/vzj2006.0055, 2007.
- Vereecken, H., Huisman, J. A., Bogena, H., Vanderborght, J., Vrugt, J. A., and Hopmans, J. W.: On the value of soil moisture measurements in vadose zone hydrology: A review, *Water Resour. Res.*, 44, W00D06, doi:10.1029/2008WR006829, 2008.
- Vogel, T., Huang, K., Zhang, R., and van Genuchten, M. T.: The HYDRUS code for simulating one-dimensional water flow, solute transport and heat movement in variably-saturated media, *Research Report No. 140*, US Salinity Laboratory, Agricultural Research Service, USDA, Riverside, CA, 1996.
- Vrugt, J. A., Bouten, W., and Weerts, A. H.: Information content of data for identifying soil hydraulic parameters from outflow experiments, *Soil Sci. Soc. Am. J.*, 65, 19–27, 2001.
- Vrugt, J. A., Bouten, W., Gupta, H. V., and Hopmans, J. W.: Toward improved identifiability of soil hydraulic parameters: On the selection of a suitable parametric model, *Vadose Zone J.*, 2, 98–113, 2003.
- Walker, J. P.: Estimating soil moisture profile dynamics from near-surface soil moisture measurements and standard meteorological data, *Ph D. thesis*, The University of Newcastle, Callaghan, New South Wales, Australia, 766 pp., 1999.

- Walker, J. P., Willgoose, G. R., and Kalma, J. D.: One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms, *Adv. Water Resour.*, 24, 631–650, 2001.
- Weerts, A. H. and El Serafy, G. Y.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, *Water Resour. Res.*, 42, 1–17, doi:10.1029/2005WR004093, 2006.
- Xie, X. and Zhang, D.: Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter, *Adv. Water Resour.*, 33, 678–690, 2010.